

AD-A203 111

Naval Medical Research Institute

Bethesda, MD 20814-5066

NMRI-24

November 1988



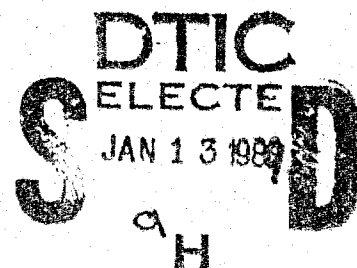
RECONSTRUCTION OF 3-D POSITRON EMISSION WITH MAXIMUM LIKELIHOOD

PAUL K. WEATHERSBY,
SHALINI S. SURVANSKI,
and PAUL MEYER

Approved for public release
distribution is unlimited

Naval Medical Research
and Development Command
Bethesda, Maryland 20814-5044

Department of the Navy
Naval Medical Command
Washington, D.C. 20372-5210



89

1 13 064

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION Unclassified			1b. RESTRICTIVE MARKINGS		
2a. SECURITY CLASSIFICATION AUTHORITY			3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution is unlimited		
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE			5. MONITORING ORGANIZATION REPORT NUMBER(S)		
4. PERFORMING ORGANIZATION REPORT NUMBER(S) NMRI 88-24			7a. NAME OF MONITORING ORGANIZATION Naval Medical Command		
6a. NAME OF PERFORMING ORGANIZATION Naval Medical Research		6b. OFFICE SYMBOL (If applicable)	7b. ADDRESS (City, State, and ZIP Code) Department of the Navy Washington, D.C. 20372-5120		
6c. ADDRESS (City, State, and ZIP Code) Bethesda, Maryland 20814-5055		8a. NAME OF FUNDING/SPONSORING ORGANIZATION Naval Medical Research and Development Command			
8b. OFFICE SYMBOL (If applicable)		9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER			
8c. ADDRESS (City, State, and ZIP Code) Bethesda, Maryland 20814-5055		10. SOURCE OF FUNDING NUMBERS			
		PROGRAM ELEMENT NO. 63713N	PROJECT NO. M0099	TASK NO. 01A-1002	WORK UNIT ACCESSION NO. DN177792
11. TITLE (Include Security Classification) Reconstruction of 3-D positron emission with maximum likelihood					
12. PERSONAL AUTHOR(S) Weathersby PK, Survanshi SS, Meyer P					
13a. TYPE OF REPORT Technical report		13b. TIME COVERED FROM TO		14. DATE OF REPORT (Year, Month, Day) Nov. 1988	
15. PAGE COUNT 36					
16. SUPPLEMENTARY NOTATION					
17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)		
FIELD	GROUP	SUB-GROUP	image maximum likelihood		
			positron planar detectors		
			tomography		
19. ABSTRACT (Continue on reverse if necessary and identify by block number)					
20. DISTRIBUTION/AVAILABILITY OF ABSTRACT <input checked="" type="checkbox"/> UNCLASSIFIED/UNLIMITED <input type="checkbox"/> SAME AS RPT. <input type="checkbox"/> DTIC USERS					
21. ABSTRACT SECURITY CLASSIFICATION Unclassified			22a. NAME OF RESPONSIBLE INDIVIDUAL Phyllis Blum, Information Services Division		
22b. TELEPHONE (Include Area Code) 202-295-2188			22c. OFFICE SYMBOL ISD/ADMIN/NMRI		

DD FORM 1473, 84 MAR

83 APR edition may be used until exhausted.
All other editions are obsolete.SECURITY CLASSIFICATION OF THIS PAGE
UNCLASSIFIED

CLASS - CVC - 5 PAGE

1

TABLE OF CONTENTS

	Page
ABSTRACT	1
ACKNOWLEDGEMENTS	iv
INTRODUCTION	1
Reconstruction Geometry and Algorithm	2
Simulation Procedure	7
SIMULATIONS AND RESULTS	10
Simple Object	10
Complex Object Simulation	18
DISCUSSION	29
REFERENCES	33

LIST OF FIGURES

Figure 1 Simple Object Sketch.	11
Figure 2 Simple Object Images.	12
Figure 3 Simple Reconstruction Statistics.	14
Figure 4 Reconstruction of Simple Module H	16
Figure 5 Complex Object Images	20
Figure 6 Complex Reconstruction Statistics	22
Figure 7 Recovery Within Complex Boundary.	25
Figure 8 Recovery Within Organ-sized Regions	27
Figure 9 Graininess Within Organ-sized Regions	28



Date and Unannounced Justification		<input checked="checked" type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
By _____		
Distribution/ _____		
Availability Codes		
Dist	Avail and/or Special	
A-1		

ACKNOWLEDGEMENTS

The authors are very grateful to Mr. Harry Fiske for excellent graphics, and Ms. S. Cecire and Ms. J. Gaines for editorial assistance. Supported by Naval Medical Research and Development Command Work Unit 63713N-M0099.01A.1002.

The opinions and assertions contained herein are the private ones of the authors and are not to be construed as official or as reflecting the views of the Department of the Navy or the Department of Defense.

INTRODUCTION

Positron emission tomography (PET) is a growing technique for medical diagnosis. Special purpose machines have been developed that achieve good resolution in two-dimensional (2-D) slices of the head or abdomen with relatively intense radioactive sources. The emphasis on these specialized 2-D devices seems to have been driven by computational advantages of reconstruction as much as by medical need. However, the original PET geometry of 2 planar Anger cameras (1) continues to have application if the emitter object is large or irregular in shape, or otherwise fails to fit the 2-D devices. Reconstruction of full 3-D images has been reported by many authors (2-8). However, these methods have limited ability to deal with low counts (6), or to use the full data set (2,3,8).

Recently, a 2-D PET reconstruction method has been described by Shepp and Vardi that uses maximum likelihood (ML) (9-11). That general statistical approach allows one to use all available physical knowledge to reconstruct an image that has the highest probability of generating the actual data set. Attempting such an optimization is a formidable process since each of the thousands of volume element intensities is an unknown random variable that co-varies with each other. The expectation-maximization (E-M) (11,12) algorithm for likelihood maximization has been shown to possess many desirable statistical and practical aspects that allow approach to ML estimates in reconstruction tomography (9-15). The problems with Fourier inversion artifacts such as negative estimated activity (16) is avoided, and all available data can be used. In 2-D devices, the ML formalism has been extended to efficient recovery of regions of interest (17) and to estimation of kinetics from a time sequence of images (14,15).

Here we report application of the ML approach to the full 3-D problem. We were motivated by a study attempting to estimate the distribution and kinetics of $^{13}\text{N}_2$ gas in human divers (related to mechanisms of decompression sickness) over time spans of many isotope half-lives. Thus, our application had vanishingly low count rates from a complex emitter with high activity in areas not of interest (18). In this paper we first pose the general problem, specify the reconstruction procedure, and then examine the algorithm performance with a series of simulated images. These images include a simple set of large rectangular solid modules, a complex emitter similar to a human experiment (18), and the complex emitter with added instrument noise (19). Performance is assessed by several statistical measures as well as by image appearance. A brief comparison is made to an algorithm proposed by Lim et al. (5,20), which is called iterative Weighted Backprojection (WB).

Reconstruction Geometry and Algorithm

The acquisition device is a pair of large (40 cm diam), stationary NaI crystals with parallel planar surfaces outside the emitting source, and connected in coincidence (21). Photons satisfying an energy and coincidence criterion trigger an A/D converter to sample and store a pair of coordinates from both the A and B camera detectors (coordinates called X_a , X_b , Y_a , Y_b). Calibration of the device is a problem discussed elsewhere (19), but calibration and camera performance parameters can be included in a ML algorithm; below, we show that they can seriously affect an image.

To organize the problem by the Shepp-Vardi approach, we first discretize the original data according to which X-Y area element on each camera face the photons arrive from. In practice, we use areas 1.33 cm on a side (a 32 x 32 square grid slightly circumscribes the circular area.) Although there is some controversy about the effect of early discretization (16,28), the limited

resolution seems to justify this step. (The actual spatial resolution of any detector depends on counting statistics, and our choice of resolution elements should be considered essentially arbitrary). Each combination of an area on one camera face with an area on the other is termed a detector "tube", d , in order to use the Shepp-Vardi nomenclature. In 2-D applications, each of these tubes correspond to a physical detector. The total number of such tubes, D , is quite large: $32^4 \sim 10^6$ in our application, so most tubes actually have zero events.

The "image" is also discretized into a number of boxes, b , set on a rectangular grid. We use the same X-Y grid as for the detector tubes and a Z grid of 8 boxes deep between camera faces. With the 45.6 camera face separation in our experiments, the Z-direction boxes are then $45.6/8 = 5.7$ cm high. This choice acknowledges the relatively poor Z-direction resolution that is inherent in the device that does not sample events emitted at a large angle from the camera axis (1). The total number of boxes in the image, B , is therefore $32 \times 32 \times 8$ or about 8000. (In practice, we place the cylindrical imaging space image within the square array so only 6,038 boxes can be part of the image).

The reconstruction problem can now be posed: estimate the B random variables $L(b)$, the emitter density in each box, given the data set $n(d)$, the recorded number of events in each detector tube. The maximum likelihood estimate of this problem is the set of L 's that maximize the overall probability of achieving the actual recorded data set. Derivation and properties of the mathematical problem are not presented here; the reader is referred to the excellent presentations of Shepp and Vardi (9-11) and of Lange and Carson (12). The likelihood function, f , of the data, n , given the current emission parameter estimates, L , can be constructed by using the

Poisson distribution on emissions from each box (10):

$$f(n|L) = \sum_{d=1}^D [-L(d) + n(d) \cdot \ln\{L(d)\} - \ln\{n(d)!\}] \quad [1]$$

where:

$$L(d) = \sum_{b=1}^B L(b) \cdot p(b,d)$$

The final term in Eqn. [1] does not depend on the estimated emission parameters, L , and can therefore be ignored. Each term $p(b,d)$ is the probability that a positron emission from image box, b , will be recorded in detector, d . This matrix incorporates aspects of both physics and detector performance, which can change in different applications. Shepp and Vardi applied the expectation-maximization (E-M) approach to achieving a ML estimate (9-12), and obtained the following algorithm:

$$L(b)_{\text{new}} = L(b)_{\text{old}} \cdot \sum_{d=1}^D \frac{n(d) \cdot p(b,d)}{\sum_{b'=1}^{B'} L(b')_{\text{old}} \cdot p(b',d)} \quad [2]$$

where $b = 1, 2, 3, \dots, B$

The outer summation in Eqn. [2] is over all possible tubes; the inner sum in the denominator is over all the boxes, B' , which have a finite $p(b,d)$ for the specific tube, d .

Since this is an iterative procedure, an initial set of $L(b)$ must be provided. The E-M algorithm has been proven as monotonically convergent (11,12), so in principle any nonzero starting image will eventually lead to the ML image. For convenience, we chose a value for each box equal to the average intensity of the entire data set (i.e., a uniform gray image). Some increase in the rate of convergence appears possible by starting with a better initial image (16), such as from a backprojection procedure.

The large $p(b,d)$ matrix is the key item that incorporates the physical and instrumental characteristics of a given detector and emission source.

Each term is actually a compound probability:

$$p(b,d) = p_1(b,d) \cdot p_2(d) / p(b) \quad [3]$$

where:

$p_1(b,d) = p(\text{emission from } b \text{ enters detector tube } d)$

$p_2(d) = p(\text{tube } d \text{ is recorded})$

$p(b) = \text{total probability that emission from box } b \text{ is recorded}$
in any detector

The second factor in Eqn. [3], p_2 , is a solid angle consideration depending only on the tube itself, which in this geometry is proportional to $\cos^3(\alpha)$, where α is the angle between the tube direction and the camera axis (line connecting the centers of the 2 detectors). A further correction is used for tubes that fall along the circular perimeter of the camera face: the 2nd factor is multiplied by the fraction of the tube area contained within the detector. This sharp edge cutoff seems to be responsible for some of the edge artifacts seen in the images. Point-to-point non-uniformities across the camera face (determined by flood source images) could also be included in $p_2(d)$ at an increased computational cost. The denominator in Eqn. [3] is the overall probability that an emission from box b is detected anywhere on the detector, and is proportional to the camera spatial efficiency (22).

We have used two different choices for the first term, p_1 , of the $p(b,d)$ matrix elements. The simpler $p(b,d)$, "Clean $p(b,d)$ ", uses the volume intersection of box b with tube d similar to the original approach of Shepp and Vardi (9,10). Such a choice uses assumptions of uniform emitter density in each box, no scattering of photons, no attenuation, and a perfect detector response.

For our more complex $p(b,d)$, some of the camera performance degradation already reported (19) are included. Both photon scattering and digitization problems have been found to contribute to a blurring of a point source image even when the emission plane is known a priori. Specifically, the distribution of events in the emission plane (established by simple backprojection) is well-described by a normal distribution superimposed on a low-level uniform density (19). Typical parameters are a standard deviation of 2 cm and a uniform density of 15% of the events. Therefore, our second $p(b,d)$ allows events in tube d to have arisen from many surrounding boxes located up to 3.5 box units away from the tube axis. The weighting of p_1 is proportional to the normal + uniform distribution cut to the distance where the distribution has fallen to 3% of its peak. This approach is clearly only an approximation to the actual physics because not all, probably not even most, of the spatial degradation happens in the plane of emission. However, it was computationally possible, therefore allowing a calibration procedure (19) and the image reconstruction to be performed with internal consistency. We will refer to this as the "Fuzzy $p(b,d)$."

The reconstruction is computationally intensive. The total $p(b,d)$ matrix allows over 10^8 entries, which we found impractical for storage, though using the full matrix has been explored by others (13). For each data set, we first rearranged events in a structured order by tube direction. This allowed us to only examine tubes with nonzero activities that are slightly less than the total events in the data set. Then at each iteration of Eqn. [2], first the term $p_2(d)$, then the $p_1(b,d)$ in Eqn. [3] was calculated for each family (same angle α) of tubes. The denominator, $p(d)$ in Eqn. [3], was calculated only once by summing the numerator of Eqn. [3] for each box over all tube directions, then retrieved as necessary from a storage array.

We also compared some reconstruction with the iterative Weighted Backprojection method (WB) proposed by Lim et al. (5,20). In the WB method, each event is partitioned into all boxes intersecting the detection tube with a weighting proportional to the box activity $L(b)$ estimated on the previous iteration. Unlike most Fourier transform methods, all data can be used and the estimated $L(b)$ are intrinsically non-negative. As in the original work (5,20), we used simple planar intersection rules (at the Z-direction mid-points of the same tomographic planes) instead of more complex rules for box-tube intersection. In all cases, we applied a camera spatial efficiency correction between each iteration.

Simulation Procedure

Spatial locations in all three dimensions for the simulation were in the precision of our camera's ADC unit: 0-127 units \sim 0.3 cm/unit. Each defined positron emitter object was rectangular with all boundaries matched to edges within the 32 x 32 x 8 reconstruction matrix. The simulation process used a hierarchy similar to our view of the emission physics. Each rectangular object volume was multiplied by its emission density (an integer from 1 to 60) in order to obtain a scalar proportional to the chance of an emission actually originating in that specific rectangular object. The first randomization used this scalar to choose which rectangular object the current emission came from. The next 3 random numbers established the X,Y,Z location within the rectangular object for the simulated emission. A spherically uniform direction was generated to simulate emission of the annihilation photons. The intersection of this direction with each camera face was calculated to see whether the event was "recorded" by the camera. The fraction of emissions recorded is proportional to the overall camera spatial efficiency for this

rectangular box, $p(b)$ (22). The simulated data from this process we call the "Clean Sim."

To deal with camera calibration performance and scattering in the simulation, two additional features were added to cause the simulated event to be placed in the "wrong" tube. A Gaussian error of specified SD in a random direction was applied at each camera face for every intersecting event. In addition, every 7th event was placed at a uniformly random location within the camera A/D range to simulate the "white noise" component of camera performance (19). These data were termed the "Fuzzy Sim."

A DEC pseudorandom number generator giving a uniformly distributed variable was used in all these processes (23). Standard transformations (24) were applied to get uniformly distributed angles and normal deviates when needed. The process was checked by simulating some point sources and comparing detection efficiency to an analytical expression for efficiency (22). The simulations required about 16 min per 10,000 detected events on a PDP 11/70. Programs were set to stop upon arrival at a predetermined total number of detected events (from 2,000 to 100,000). An auxiliary output file recorded the total events actually emitted from each rectangular box, $S(b)$. This simulated image could then be used to assess the performance of reconstruction. Because of the high computational cost, simulations were generally not repeated.

In addition to likelihood itself, we assessed image recovery by a root mean square error (rms) and a weighted root mean square error (wrms) based on simple or weighted sums of squared deviations. These measures require a known image for comparison and thus are useful only for simulation, not for unknown objects.

$$\text{rms} = \sum_{b=1}^B \frac{1}{B} \left([L(b) - S(b)]^2 \right)^{1/2} \quad [4]$$

$$\text{wrms} = \frac{1}{B} \left(\sum_{b=1}^B \frac{[L(b) - S(b)]^2}{[\text{Max}(S(b), 1)]} \right)^{1/2} \quad [5]$$

As before, the L-array contains the recovered emission density, while the S-array has the actual simulated density. The rms is an average deviation from the simulated image in counts per box. The wrms is an attempt to account for one known major source of uncertainty in a real image: Poisson counting statistics. For each independent Poisson process, such as the emissions from one box that is detected at all, the expected precision (standard deviation) of that count can be used for weighting error summaries. For large counts, Poisson standard deviation is the square root of the raw count. (In Eqn. [5] the denominator is the square of this expected standard deviation.) We realize that for a simulation the use of this formula is not precise, both because of low counts where the square root is not a good approximation, and because we know S(b) exactly after the Poisson uncertainties of emission and detection have occurred. Subject to these limitations, the wrms is used as an average dimensionless error per box, where wrms = 1.0 would be image recovery to the expected limit of counting statistics. A wrms error much larger than 1.0 means more variability than in the simulation. A wrms error much smaller than 1.0 means the algorithm is "too precise" in attempting the reconstruction. Image recovery more precise than the original simulation has recently been noted with the E-M algorithm in some 2-D reconstructions (28,29).

SIMULATIONS and RESULTS

Simple Object

The objective for using this first object was examination of how the reconstruction algorithm would recover object boundaries, especially when using small total count numbers. A secondary objective was development of a battery of test statistics to describe image recovery. Figure 1 is a sketch of the simple object simulated. The object was a set of 10 large modules, 9 to 91 boxes each; adjacent modules were simulated to have intensity contrast differences of about 2 going in both X and Z directions. In addition, a large low-intensity object (module G in Fig. 1) was used to add out-of-focus events that would complicate the reconstruction of modules in the next Z-plane (modules H,I,J). No part of the modules extended into the Z-planes closest to each camera face (levels 1 or 8). Three sets of simulations were done with 2000 to 50,000 total counts to span the count range of 4 to 659 counts per individual box. Simulation was "clean", that is without inclusion of the camera degradation. The reconstruction only used the first of the $p(b,d)$ methods described above (Clean Recn: no error from camera).

Performance of both the ML and WB reconstruction algorithms for this simple object are shown in the next several figures. All 8 levels of the original 50,000-count simulated object and the 50th iteration by both reconstruction methods are shown in Fig 2. All images have been corrected for camera efficiency by dividing each box count by $p(b)$ such that the center boxes have the actual count at each level, and boxes extending out radially have a progressively exaggerated count. The edges are successfully recovered, but with less blurring for the ML procedure than for WB. Very few of the counts are distributed outside the boundaries of the original objects. Any counts assigned to levels 1 and 8 are erroneous as no source locations were

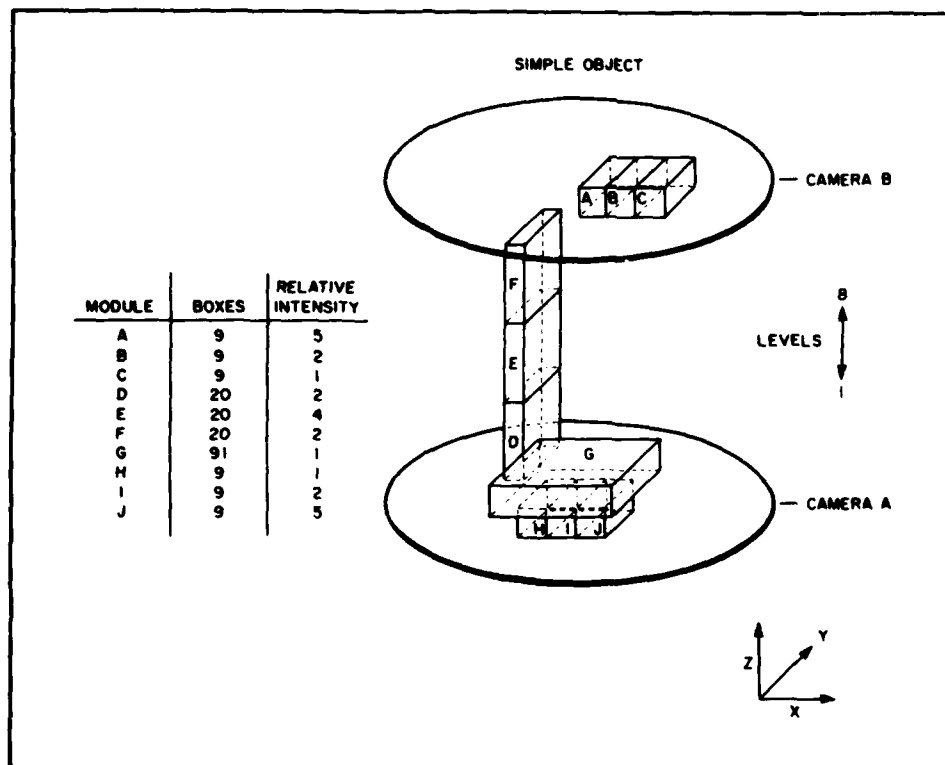


Figure 1. Simple Object Sketch. The 10 large modules (9 to 91 boxes each) were simulated to have intensity contrast differences of about 2 going in both X and Z directions. The large low intensity object (module G) was used to add out-of-focus events to complicate the reconstruction of modules in the next Z plane (modules H,I,J). No part of the modules extended into the Z planes closest to each camera face (levels 1 or 8).

RECOVERY OF SIMPLE OBJECT

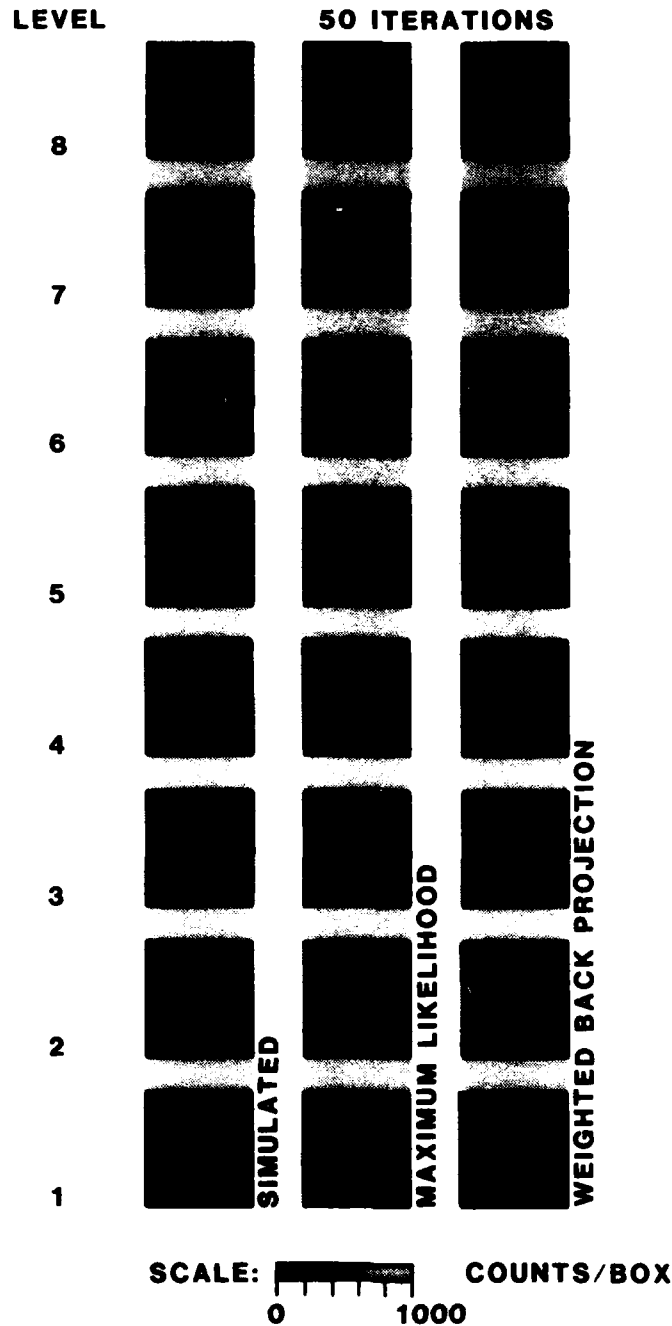


Figure 2. Simple Object Images. Gray scale images of the simple object simulated with 50,000 total counts and the recovered images after 50 iteration by both maximum likelihood and weighted back projection methods. All images have been corrected by the camera spatial efficiency matrix, $p(b)$, such that only boxes in the center of each level have the correct raw count intensity.

simulated there. Recovery of detail is slightly better in mid levels than close to either camera face. Note also a bit of increased "graininess"; the individual boxes have more apparent variability about their mean intensity than was simulated in the original object. The occasional speckle in the outer corners of the images is a reconstructed box density that has been amplified by the efficiency correction whose magnitude is greatest in the corners.

Recovery with successive iterations are presented next. As shown in the top of Fig. 3, overall likelihood improved substantially for the first 20 iterations of the ML algorithm, Eqn. [2], and only slowly thereafter. Some finite improvement was noted even after 200 iterations. Log likelihood differences less than order 1 are not important in statistical applications with few parameters, so image improvement after 20 to 50 iterations appears marginal. The likelihood stabilizes with little subsequent improvement in later iterations with smaller data sets. Thus, an arbitrary rule for declaring the "convergence" of the final image would be needed. This stopping rule was examined for all the simulated reconstructions. The WB method produced images that initially improved as measured by likelihood but then decreased in likelihood. The decrease occurred on later iterations in data sets with higher total number of events in the image.

In the bottom of Fig. 3 are plots of both rms and a wrms error for the simple object. The figure shows that these criteria improve (i.e., decrease) with subsequent iterations using ML, but very slowly after 20 iterations. Curves that show an initial rise indicate that early iterations produce images with poorer statistics than the original, uniformly gray initialization. Weighted rms error was always less than raw rms, and eventually decreased to about a factor of 3 higher than the "perfect" value for a Poisson process. In

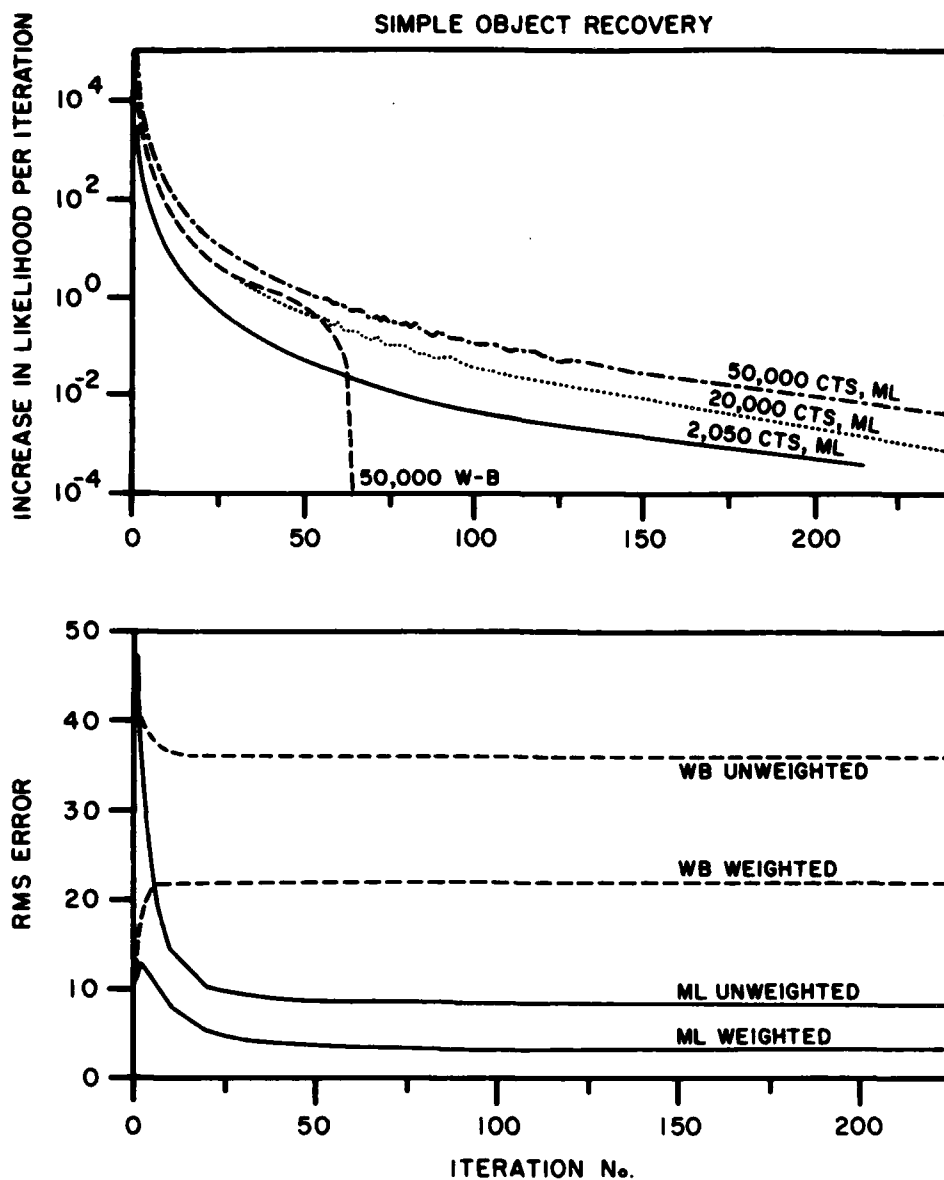


Figure 3. Simple Reconstruction Statistics. Overall image recovery statistics for the simple image. Both increase in value of the likelihood function, f , and changes in mean-square error are plotted against increasing iterations of the reconstruction algorithms.

other cases with fewer total events, we noted that the rms and wrms errors passed through a minimum and then increased with subsequent iterations. This means that the image originally got better by all criteria, then the image improved by ML while becoming worse by the rms and wrms criteria. The WB method stabilized within 15 iterations, but had more than a four-fold poorer performance than ML after image convergence.

For this image, all reconstructions were rather successful in avoiding assignment of events to the planes (simulated as empty) nearest the camera face. For example, at 30 iterations less than 3% of the ML reconstructed counts and 5-6% of the WB counts fell in the empty planes. Subsequent iterations decreased the fraction even lower.

Individual object recovery and the graininess problem were examined in more detail. Module H in Fig. 1 was chosen as the example in Fig. 4. This level may be most susceptible to corruption because of the large distributed object immediately above it. The upper half of the figure shows recovery of average counts for the 9 boxes in that module. Using ML, only 50% of the original counts are recovered for the lowest count rate simulated, but over 95% recovery is achieved for higher count rates. Recovery of counts using WB was poorer in all cases. (Both reconstruction methods conserve the total number of counts, so the "missing" counts from recoveries less than 100% are assigned to other boxes, most of which were simulated as empty.) The bottom half of the figure is a measure of "graininess", obtained from the standard deviation of $L(b)$ over the module. The standard deviation of reconstructed counts in the module has been normalized by the standard deviation in the original simulation so numbers greater than 1 represent more box-to-box variability or "graininess" than in the original module. By 10 iterations, the graininess had exceeded the variability in the simulated module. The

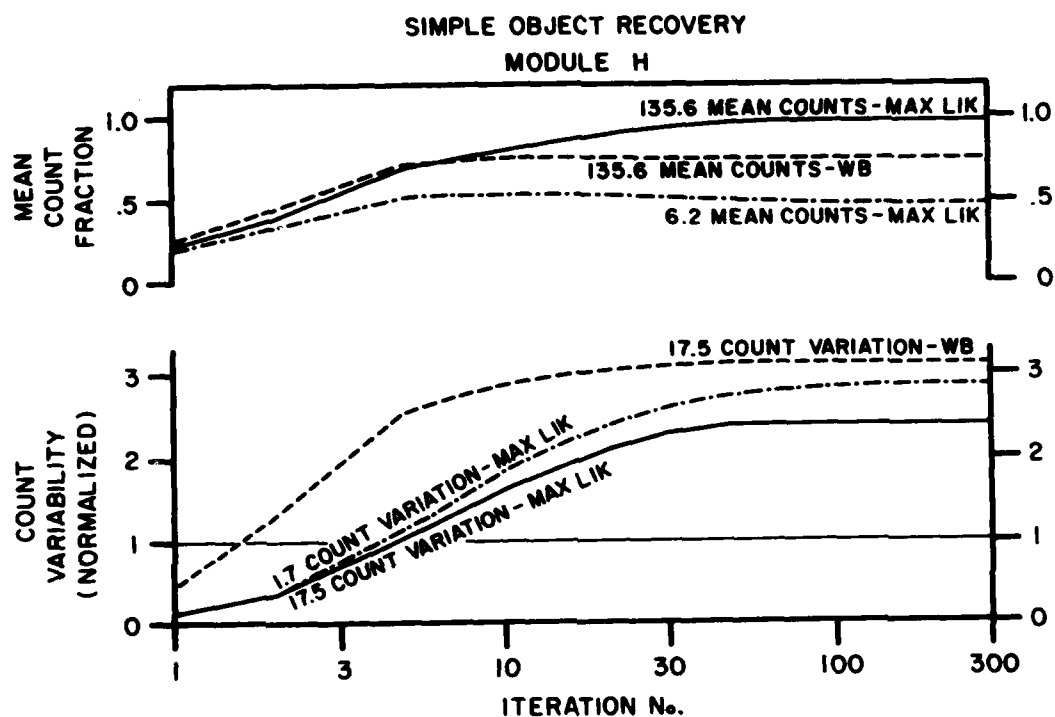


Figure 4. Reconstruction of Simple Module H. Overall count recovery and count variability for Module H of the simple image. Variability defined as the standard deviation of counts within the module at each point in the reconstruction, divided by the standard deviation of the original simulated count distribution $S(b)$.

images got progressively more grainy for the next 50 or so iterations, even though the likelihood function was showing a slow improvement (Fig. 2, Top). If reconstruction was stopped at the point of original graininess, then other desirable features of the reconstruction would be sub-optimal. This feature has been recognized as a drawback to ML, perhaps intrinsic to the method (15,28), and other recent work addresses means to avoid the problem (25,29). Our approach, like others, is to stop the reconstruction at a finite number of iterations. Other areas (not shown) had the same general result: both recovery of counts and graininess increased with successive iterations. Low count rate modules had poorer recovery and increased graininess.

An examination of the recovery of counts in the 170 individual boxes for each simulation also supported the impressions just described. For low intensity emission, average recovery is low (average 28% recovery for boxes emitting 3-20 counts); for intermediate activity recovery is better (average 66% for 21-100 counts); for high activity average recovery is excellent (average over 90% for 100+ counts). We also compared the distribution of recovery and compared it to that expected for a Poisson noise-limited process. Specifically, we calculated bounds of double the Poisson emission error on $S(b)$ with the expectation that approximately 90% of the boxes should be recovered within that band if the reconstruction itself added no image noise. Recovery within that "90% band" did not seem to depend on emission density, and overall only about half of the boxes were recovered within the band. Thus, the image noise has a component greater than intrinsic Poisson noise.

Computation times were quite different for the two algorithms. The WB method required about 4 min per 10,000 counts per iteration, while the ML method took 55 min for the same task on a PDP 11/70. Using the Fuzzy $p(b,d)$

was another three-fold slower and required us to use a CRAY XMP-2 to complete the project.

To summarize, initial exploration of the ML algorithm with this simple object was a success. The intensities in the image boxes were recovered within 50% or better despite the low count rates, and the statistical properties of the reconstruction were considerably better than the WB algorithm.

Complex Object Simulation

The other mathematical phantom was considerably more complex. For our application in experimental physiology (18), we needed to simulate a portion of a human body with a shoulder, parts of a head and trunk, and a hose that delivered radioactive nitrogen gas to a mouthpiece. The purpose was to study gas delivery to non-gas tissues, such as the shoulder joint implicated in diver decompression sickness. The mathematical phantom for this case used 37 rectangular modules of various sizes. As in the simple object, no emissions occurred in the Z-levels immediately adjacent to the camera face. Since nitrogen solubility in human tissue is rather low (26), the simulated hose and mouth were set to emit at a sixty-fold higher intensity than tissue; and the simulated lungs and trachea at 40 times the tissue level. Together these gas-filled areas accounted for slightly over 90% of the total simulated emissions.

The objectives here were to apply the insights and measures developed above to build a reconstruction scheme useful in our difficult application: complex emitter, low counts, data dominated by presence of non-interesting objects, and camera performance degradation. We present 3 cases of about 50,000 events: simulation without camera degradation and $p(b,d)$ defined as tube-box volume intersection ("Clean Sim, Clean Recn"); simulated data with camera degradation, but with the simple $p(b,d)$ ("Fuzzy Sim, Clean Recn"); and

the same degraded data but with reconstruction using the less well localized $p(b,d)$ ("Fuzzy Sim, Fuzzy Recn"). These data were not examined with the WB algorithm.

The original object and the reconstructed images are shown in Fig. 5. For orientation, the view is of a supine human. The gas delivery hose goes across level 7 and connects to a mouthpiece in levels 5 and 6. The lower right sections of the subject's head is in the upper right quadrant of levels 3-5; his trunk to the right shoulder covers the lower half of levels 2-4. In the simulation, no activity is simulated in levels 1 or 8. The highest activity corresponds to the simulated mouth and hose in levels 5-7. Nearly as high activity is in the simulated subjects' lungs and airways in levels 3 and 4. The object of our physiological experiment is recovery of activity (and kinetics) in the non-gas lower activity regions mostly in levels 2-4. The images reconstructed for the 3 treatments of image degradation are also in Fig. 5. All reconstructions have some degree of inappropriate assignment of activity to levels 1 and 8; all recover much of the high activity regions in levels 3-7; and all have some activity assigned to the desired regions in levels 2-4. These features will be compared in more detail below.

Other visual aspects of the images deserve comment. All reconstructions have high-activity speckles and "rings", especially in levels 1 and 8 of Fuzzy-Clean. The magnitude of the problem has been exaggerated by the image normalization procedure used, since this device has a spatial sensitivity that drops sharply along the outer circumference of the detector and more gradually toward each camera face (22). To avoid this efficiency gradient dominating the image appearance, we divide all box activities by the local efficiency; the procedure introduces a modest correction in the center of the camera but is a large correction along the outer circumference. Thus, the

RECOVERY OF COMPLEX OBJECT

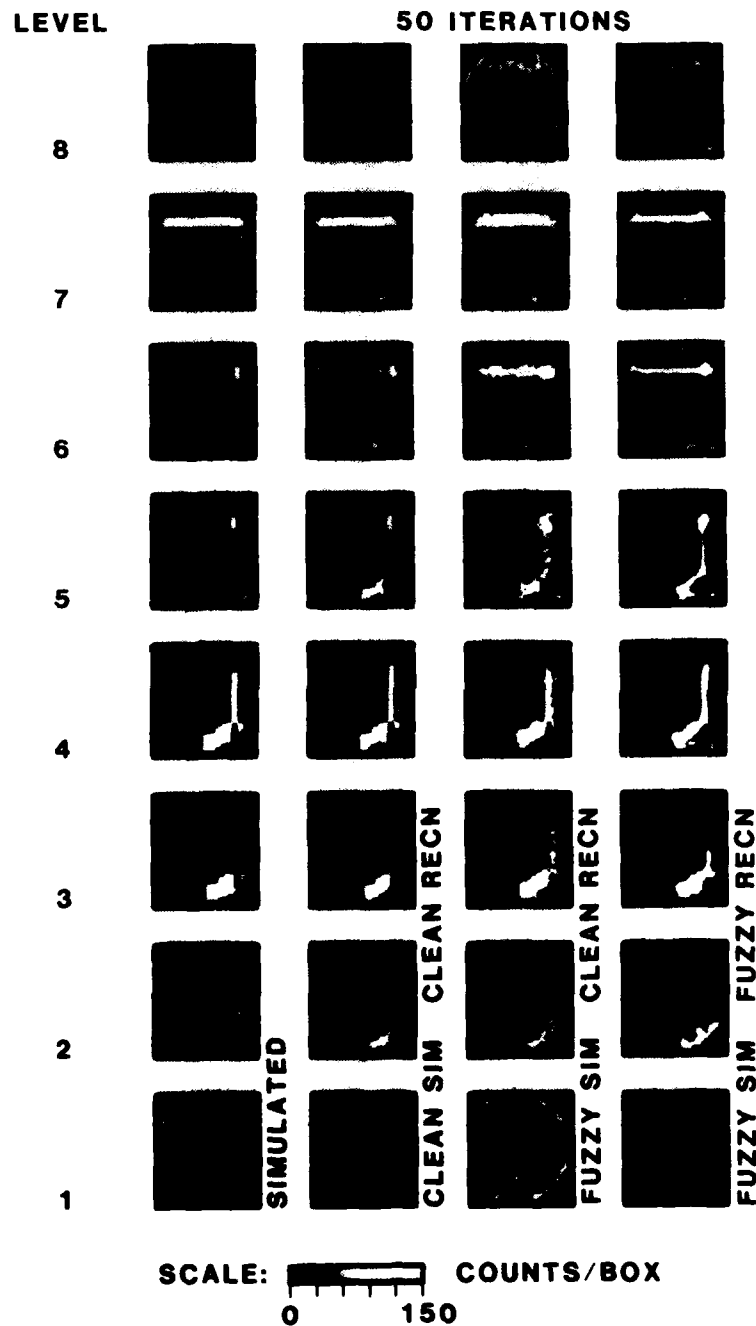


Figure 5. Complex Object Images. Gray scale images of the complex object simulated with 50,000 total counts and the recovered images after 50 iterations for the cases of Clean Simulation, Clean Reconstruction; Fuzzy Simulation, Clean Reconstruction; and Fuzzy Simulation, Fuzzy Reconstruction. All reconstructions used maximum likelihood. All images have been corrected by the camera spatial efficiency matrix, $p(h)$, such that only boxes in the center of each level have the correct raw count intensity.

information content along the outer edge is intrinsically low and activity in these regions should be ignored in quantitative interpretation despite its striking visual appearance.

Another visible problem is the anticipated poor resolution in the Z direction (along camera axis). The problem is most easily seen as the "shine" of hose activity simulated in level 7, but partially recovered in levels 6 and 8. It appears to a lesser extent in the lung regions of levels 2-5. The Fuzzy simulations are worse than the Clean Sim. We noted that the problem slowly improved with successive iterations for each case, but even 100 iterations in Fuzzy-Fuzzy did not eliminate the shine entirely.

In all images there appears to be a speckle pattern through regions that were simulated as homogeneous activity subject to Poisson noise. A major part of that is the deliberately low activity compared to many positron images that use a thousand-fold higher activity. (Note: the gray scale chosen has maximum contrast near 50 counts per box, an activity certainly subject to appreciable Poisson noise.) Although no specific smoothing operations were applied to these images, it is clear that the Fuzzy reconstruction produces a smoother image. The Fuzzy-Fuzzy appears smoother in the low count regions as well as the higher activity. However, there appears in all images to be a graininess problem that will be examined in more detail below.

Specific performance features of the reconstruction will now be examined, starting with the likelihood function itself. As in the other simulation and in other applications of the E-M algorithm, the likelihood function improved rapidly in the early iterations as seen in the top of Fig. 6. Improvement thereafter was ever more gradual. The Clean-Clean image converged to a stable likelihood fastest; that is, it had a large improvement in likelihood in the early iterations, but could only make small improvements by iteration 50.

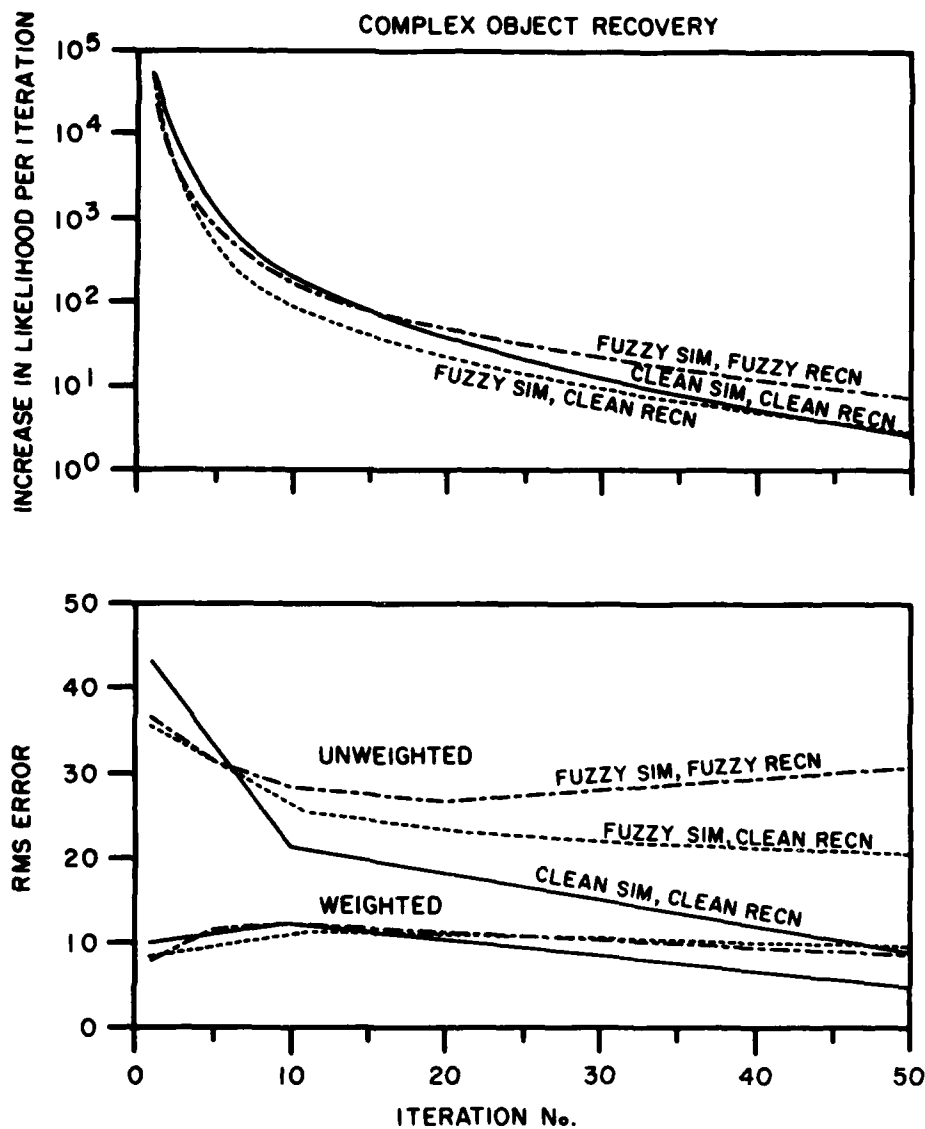


Figure 6. Complex Reconstruction Statistics. Overall image recovery statistics for the complex image. Both increase in value of the likelihood function, f , and changes in mean-square error are plotted against increasing iterations of the reconstruction algorithm.

This favorable behavior might be expected for a process without the ambiguities in source added by the Fuzzy processes. By comparison, for both fuzzy images there was a smaller improvement in likelihood initially followed by a more sustained increase of 1-10 units per iteration even after 30 iterations, thus showing a slower convergence. Since both Fuzzy-Fuzzy and Fuzzy-Clean have the same data, it is possible to compare absolute values of the likelihood function itself by direct use of Eqn. [1]. The very slowly converging Fuzzy-Fuzzy produces a likelihood poorer by 3,342 units at iteration 50. Thus, the Fuzzy-Fuzzy reconstruction is statistically poorer overall than Fuzzy-Clean at that point (this difference is 0.55 per box so we are not certain that the difference is very significant). Since the rate of Fuzzy-Fuzzy convergence is so slow, it seems possible that "eventually", when a maximum likelihood is achieved, the Fuzzy-Fuzzy might be a statistically superior image. Unfortunately, that point did not appear practically attainable, as another 50 iterations on Fuzzy-Fuzzy only attained 1/20 of the likelihood discrepancy. Overall, we have the impression that the rate of convergence decreases with the complexity of the reconstruction process. With the same amount of data, the simple object was reconstructed faster than the complex object and the Clean $p(b,d)$ converged faster than the Fuzzy $p(b,d)$.

The lower section of Fig. 6 shows the rms error criteria for the 3 cases. In all 3 cases the rms improved markedly in early iterations, while wrms increased before establishing a slow improvement. It appears that the E-M algorithm corrected major deficiencies in the high count areas before settling into a slow correction over all boxes. Again Clean-Clean had the best behavior, decreasing both the unweighted and weighted statistics with successive iterations, and achieving the lowest values of both statistics. Fuzzy-Clean improved its rms error for 10-20 iterations, then had no further

change. As with the rate of convergence by likelihood, Fuzzy-Fuzzy had the poorest performance by actually increasing its unweighted rms error after 20 iterations. The divergent trends of rms and wrms errors implies a difference due to the weighting. Specifically it appears that for the Fuzzy-Fuzzy case later iterations progressively increase graininess in high count boxes with little impact on the likelihood of the overall image.

Now we will examine the reconstruction of parts of the overall image. First an edge boundary question: what fraction of the counts are assigned somewhere within the region with non-zero emissions. Since the reconstruction preserves the total number of counts, all assignments outside this region are in error. As seen in the upper section of Fig. 7, the Clean-Clean recovery is excellent, with nearly 90% recovery by iteration 50. The Fuzzy sets were poorer with only 62% for Fuzzy-Clean and 77% for Fuzzy-Fuzzy. However, most of these counts were in areas of simulated high counts (corresponding to gas hose and subject's lungs) and thus not of physiologic interest. Just considering the important areas of non-gas tissue (not plotted), the fuzzy sets were better: both recovered over 80% while the Clean Sim had about 60%. Thus, it appears that the error-free process does a better job with the high intensity areas and a poorer job with the low count rate regions of interest. A similar observation has been made on error-free simulations in 2 dimensions (28,29). Some of this comparison is artifact, however, as the Fuzzy Sim had a substantial fraction of counts that were not assigned to any box. (For generating reconstruction statistics we only use the 84% of the total data in matrix $S(b)$).

The next concern is axial resolution: how successful was the reconstruction in proper positioning of counts along the coordinate connecting the two camera faces. In prior use of the WB method, we noticed that "hose"

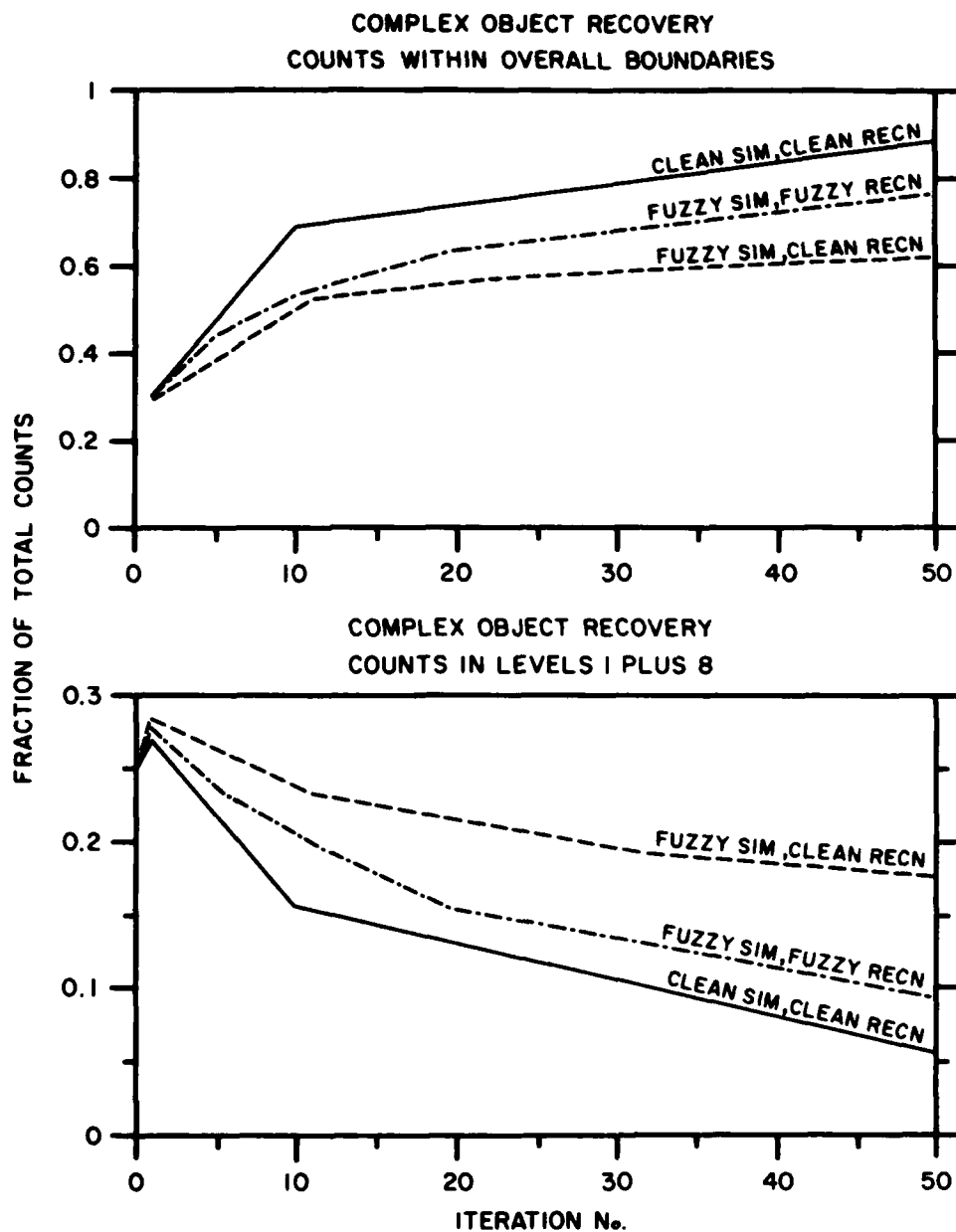


Figure 7. Recovery Within Complex Boundary. Aspects of complex image reconstruction. Upper panel shows fraction of simulated counts recovered within overall boundary of original simulated object. Lower panel shows fraction of counts assigned to the 2 levels immediately adjacent to the camera (levels 1 and 8) which had no counts simulated.

activity in level 7 was often incorrectly assigned to level 8. Figure 7 (lower panel) has the fraction of total counts assigned to levels 1 and 8, where the simulation actually had none. In all 3 cases successive iterations reduced this incorrect assignment. Best performance was again the Clean-Clean, but the Fuzzy reconstruction achieved better axial resolution of the Fuzzy data than did the Clean reconstruction. After 50 iterations, less than 10% of the total counts were left in the originally empty levels 1 and 8. Though more difficult to quantitate, other axial degradation features were not as favorable with Fuzzy-Fuzzy. Specifically, the "lung" activity in levels 3,4 seemed to be more spread into adjacent planes with Fuzzy-Fuzzy compared to Fuzzy-Clean.

Image recovery in organ-sized regions of multiple boxes varied with region. Two regions are reported in Fig 8. In the lung region (bottom, an area of 42 boxes in levels 3 and 4 with an average of 85 counts per box simulated), average recovery varied from about 60% to over 95% for the different cases, with the Fuzzy-Fuzzy being the best. Many iterations were required for that recovery, and actually the recovery was increasing even after 50 iterations. In an area of more interest, the shoulder (top, an area of 78 boxes in levels 2-4 with 5.6 counts per box simulated), recovery was 70-85% by 50 iterations. Again the Fuzzy-Fuzzy had greatest recovery, and again the recovery increased steadily through at least 50 iterations (it achieved 92% after 100 iterations). Other areas examined (but not presented in figures) had a similar outcome: recovery of 50-150% of average simulated counts, with most low-activity areas of interest recovering 80-125% and with the best performance resulting in the Fuzzy-Fuzzy case.

Even with good recovery on average, deviations around the average were a problem. Figure 9 shows the previously defined graininess for the same areas

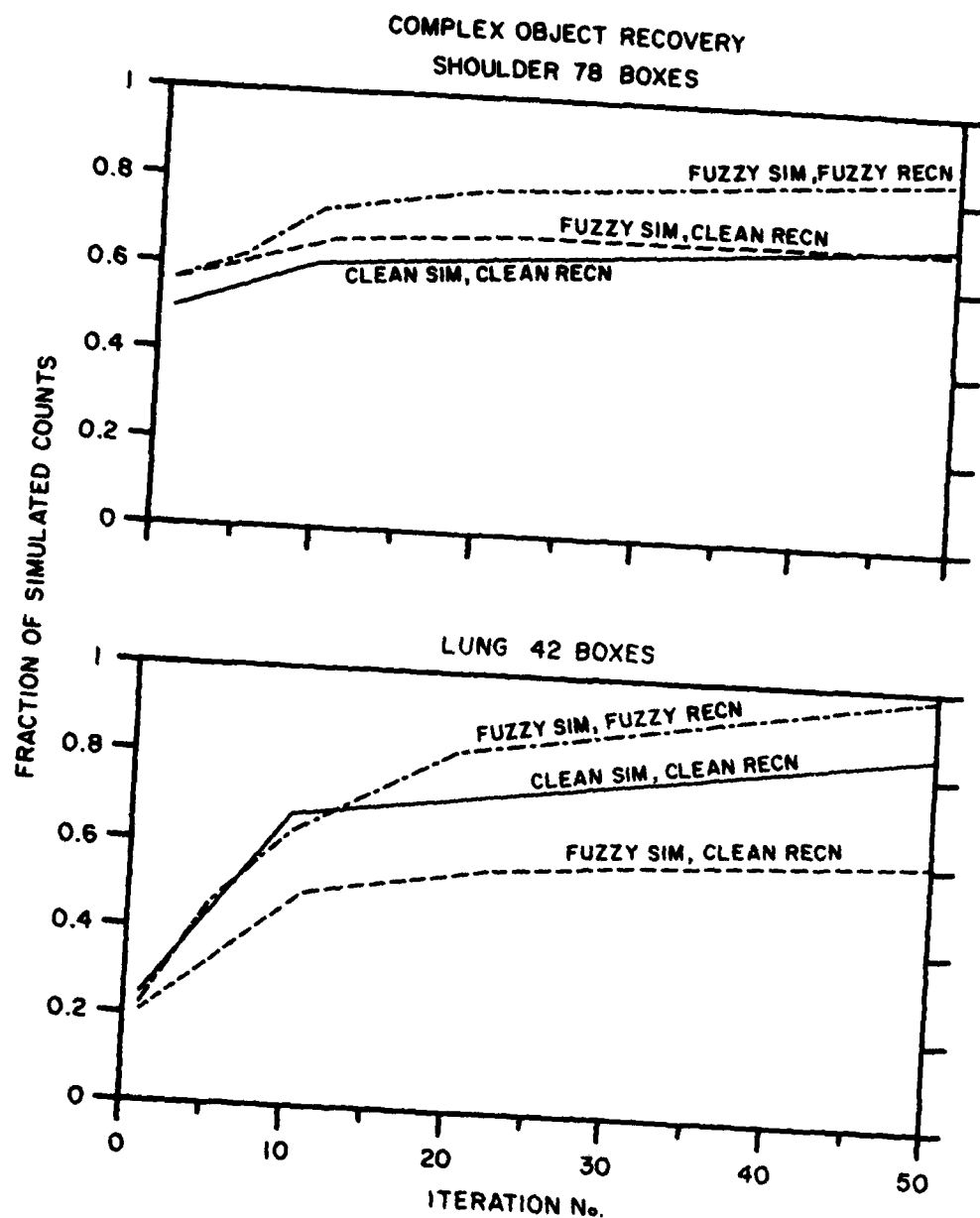


Figure 8. Recovery Within Organ-sized Regions. Recovery of counts in two areas of the complex image identified by their simulated anatomic location: shoulder (top) and lung (bottom).

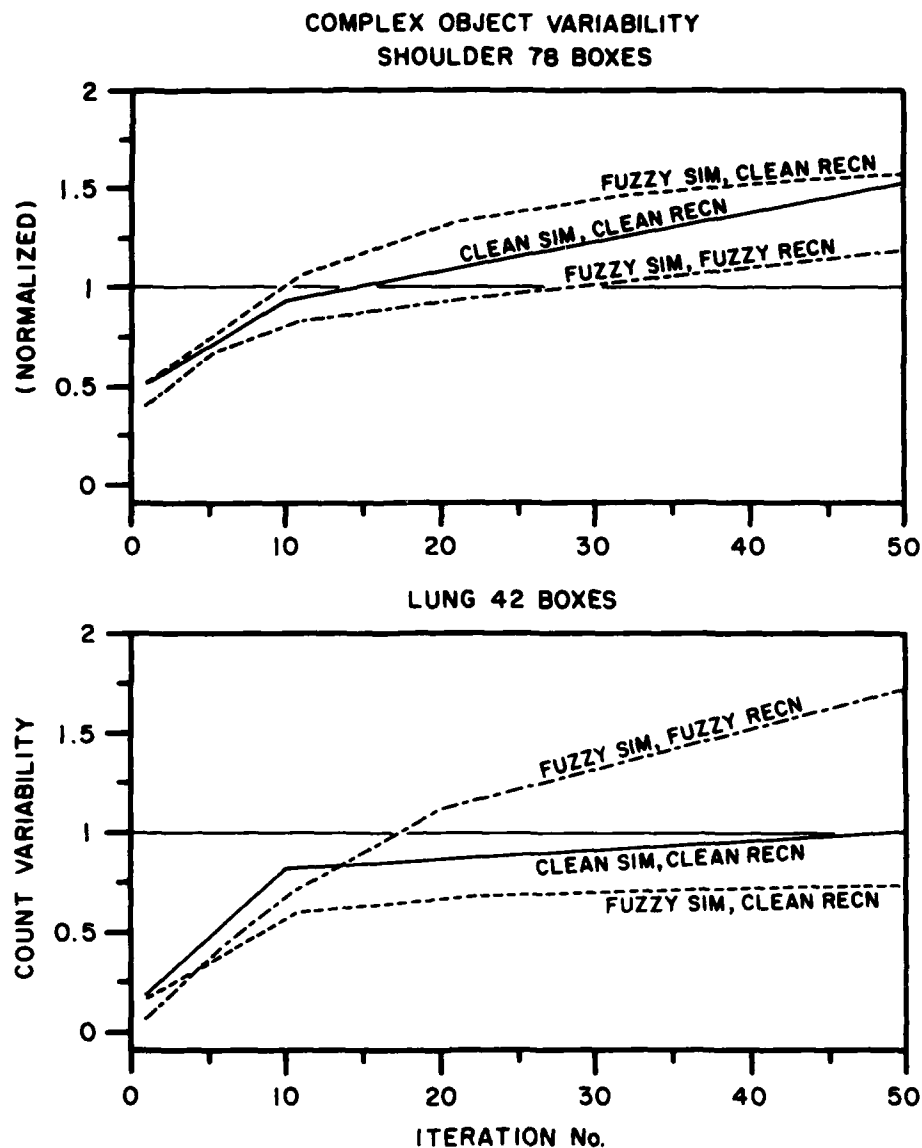


Figure 9. Graininess Within Organ-sized Regions. Overall count recovery and count variability for the two areas of the complex image as used in Fig. 8. Variability defined as the standard deviation of counts within the module at each point in the reconstruction, divided by the standard deviation of the original simulated count distribution $S(b)$.

as in Fig. 8. As was seen in the simpler object, the graininess increased progressively during the reconstruction. For the lung area, the graininess stabilized at slightly less than the original simulated level for Clean-Clean and Fuzzy-Clean, while it continued to rise for Fuzzy-Fuzzy. In the case of the shoulder, the increase was continual through 50 iterations. For that region of physiological interest, the Fuzzy-Fuzzy had barely exceeded its original level by 50 iterations. As a general observation, the high count areas had less of a graininess problem with Clean-Clean or Fuzzy-Clean reconstructions, while the problem was least in low count areas for the Fuzzy-Fuzzy image.

Finally, image recovery on the level of single image boxes was examined. As with the simple object simulation, average recoveries increased with higher emission density. Boxes with fewer than 20 counts were recovered very poorly (10% or less of the original activity was assigned to those boxes by the reconstruction algorithm) for Clean-Clean and Fuzzy-Clean, and almost as poorly (20-30%) for Fuzzy-Fuzzy. Higher activity areas did have a better average recovery, but only above 100 counts per box did the average recovery per box exceed 70%. We also examined how many boxes were recovered within the approximate Poisson noise limits previously described. Again, no clear dependence on count rate emerged, except for both Fuzzy simulations where nearly all the high count areas were outside the nominal 90% band (only 10% of boxes within band for 100+ counts). For other cases, 20-30% of the boxes were within the band. Thus, as in the simple object reconstruction, overall reconstructed noise was 3 to 4 times higher than expected from the original Poisson counting error.

DISCUSSION

This exploration of positron image reconstruction was motivated by a

specific application requiring maximum quantitative recovery of the data (18). A small amount of published information was available to answer the following question: With these experimental data and a given reconstruction, how reliable is the estimated count density in a single box, or in an average from a small number of boxes?

Most reported reconstruction work is aimed at two-dimensional ring devices. The typical application has 100+ counts per box and sufficient overall data to average many hundred counts per detector tube. We had a different imaging geometry, a low count rate so that the average counts per detector "tube" was <1 , and a structured emitter with a sixty-fold or greater difference in emitter density.

In this environment, the ML approach was very attractive. It was based on a solid statistical treatment of the data according to the known (Poisson) distribution of the original photon emission. It allowed direct use of any a priori physical, theoretical, or empirical calibration knowledge in the reconstruction process. And it provided a single goodness-of-fit statistic (the value of the likelihood function itself) to judge "convergence" to a stable "best" image.

In practice, the present implementation of the ML algorithm did not provide a single perfect reconstruction solution. From both objects simulated, we learned that a continuously increasing likelihood is not a continuously "better" image as judged by appearance, root-mean-square statistics, or measures of graininess in the image. The divergence of statistics is not particularly bothersome, since different statistical criteria often give different sets of parameters, as happens when one changes the weighting of the original data. Even after this analysis we accept that the likelihood is the single best reconstruction performance parameter.

More disturbing is the tendency of the algorithm to increase point-to-point variability (graininess) with an apparently small increase in likelihood. Even if the algorithm was not so computationally intensive, it would appear desirable to stop the process short of "convergence". A recent report on ML reconstruction of single photon images shows continued graininess out to 10,000 iterations (27). In the more modest number of iterations attempted here, image recovery is substantially complete by a few tens of iterations. The use of the Fuzzy Recn here introduced an empirical feature that would seem to encourage smoothing rather than graininess, and there was some indications from the reconstructions that it did just that. At least, it delayed the increase in graininess to higher iterations than we were able to explore. At this point we still have no alternative to an arbitrary rule for stopping the reconstruction; 50 iterations seems a reasonable choice for the studies similar to the complex object simulated.

What final expectations can we have for applying this methodology to our actual experimental data (18)? First, we should use the Fuzzy-Fuzzy procedure, since ignoring the real known image degradation will produce unnecessary image artifacts and compromise quantitative image recovery. Then, we should expect the image to be unrealistically grainy, with an appearance of larger point-to-point variation in isotope concentration than is likely to be physically present. Finally, we expect that average count recovery will be lower than the emission by approximately 50%, and that actual box concentration estimates will vary by about a three-fold higher range than the apparent Poisson counting error.

Is this a satisfactory method for our application? Probably not. The areas of physiologic interest in our experiments (18) have activities of up to 50 counts per box. With results from the simulation, we might get images of

about 25 counts and a variability of that same magnitude. Such imprecision would not allow the production of useful images where we would be interested in point to point differences in isotope concentration in the body of less than a factor of two. In trying to estimate kinetics, the total number of counts available for analysis would be effectively higher, but not by an order of magnitude. The problems with rapid isotopic decay already limit kinetic interpretations in large regions emitting thousands of counts (18).

Additional uncertainties due to device characteristics ignored by our calibration procedure (19), and due to physical effects of photon scattering and attenuation - not considered in this simulation - would increase the final uncertainty to unproductive levels.

REFERENCES

1. Brownell G.L., Correia, J.A., and Zamenhof, R.G., "Positron Instrumentation." Recent Advances in Nuclear Medicine, Vol. 5, pp. 1-49, 1978.
2. Chang, L.T., Macdonald, B., and Perez-Mendez, V., "Axial tomography and three dimensional image reconstruction." IEEE Transactions on Nuclear Science, Vol. NS23, pp. 568-572, 1976.
3. Chu, G. and Tam, K-C, "Three-dimensional imaging in the positron camera using Fourier Techniques." Physics in Medicine and Biology, Vol. 22, pp. 245-265, 1977.
4. Hoffman, E.J., Huang, S-C, Plummer, D., and Phelps, M.E., "Quantitation in positron emission computed tomography: 6. Effect of nonuniform resolution." Journal of Computer Assisted Tomography, Vol. 6, pp. 987-999, 1982.
5. Lim, C.B., Cheng, A., and Boyd, D., "A fast iterative method for 3-D positron image reconstruction." American Nuclear Society Transactions Vol. 26, pp. 154-155, 1977.
6. Macdonald, B., Lim, C.B., Perez-Mendez, V., and Tam, K.C., "A comparison of three three-dimensional reconstruction methods for large-area positron cameras." Journal of Computer Assisted Tomography, Vol. 2, pp. 642-643, 1978.
7. Ra, J.B., Lim, C.B., Cho, Z.H., Hilal, S.K., and Correll, J., "A true three-dimensional reconstruction algorithm for the spherical positron emission tomograph." Physics in Medicine and Biology, Vol. 27, pp. 37-50, 1982.
8. Tam, K-C, Perez-Mendez, V., and Macdonald, B., "3-D object reconstruction in emission and transmission tomography with limited angular input." IEEE Transactions on Nuclear Science, Vol. NS-26, pp. 2797-2805, 1979.

9. Shepp, L.A. and Vardi, Y., "Maximum likelihood reconstruction for emission tomography." IEEE Transactions on Medicine Imaging, Vol. MI-1, pp. 113-122, 1982.
10. Shepp, L.A., Vardi, Y., Ra, J.B., Hilal, S.K., and Cho, Z.H., "Maximum likelihood PET with real data." IEEE Transactions on Nuclear Science, Vol. NS31, pp. 910-913, 1984.
11. Vardi, Y., Shepp, L.A., and Kaufman, L., "A statistical model for positron emission tomography." Journal of the American Statistical Association, Vol. 80, pp. 8-20, 1985.
12. Lange, K. and Carson, R., "EM reconstruction algorithms for emission and transmission tomography." Journal of Computer Assisted Tomography, Vol. 8, pp. 306-316, 1984.
13. Llacer, J. and Meng, J.D., "Matrix-based image reconstruction methods for tomography." IEEE Transactions on Nuclear Science, Vol. NS-32, pp. 855-864, 1985.
14. Ollinger, J.M. and Snyder, D.L., "A preliminary evaluation of the use of the EM algorithm for estimating parameters in dynamic tracer studies." IEEE Transactions on Nuclear Science, Vol. NS-32, pp. 848-854, 1985.
15. Snyder, D.L., "Parameter estimation for dynamic studies in emission-tomography systems having list-mode data." IEEE Transactions on Nuclear Science, Vol. NS-31, pp. 925-931, 1984.
16. Herman, G.T., Censor, Y., Gordon, D., and Lewitt, R.M., "Comment (on Ref.#11 above) Journal of the American Statistical Association, Vol. 80, pp.22-25, 1985.
17. Carson, R.E., "A maximum likelihood method for region-of-interest evaluation in emission tomography." Journal of Computer Assisted Tomography, Vol. 10, pp. 654-663, 1986.

18. Weathersby P.K., Meyer, P., Flynn, E.T., Homer, L.D., and Survanshi S., "Nitrogen gas exchange in the human knee." Journal of Applied Physiology, Vol. 61, pp. 1534-1545, 1986.
19. Weathersby, P.K., Survanshi, S., and Meyer, P., Robust spatial calibration of a planar positron camera, Naval Medical Research Institute Technical Report, No. 88-23, Bethesda, MD, 1988.
20. Lim, C.B., Cheng, A., Boyd, D.P., and Hattner, R.S., "A 3-D iterative reconstruction method for stationary planar positron cameras." IEEE Transactions on Nuclear Science, Vol. NS-25, pp. 196-201. 1978.
21. Behrin, E., Positron/scintillation camera data acquisition and display system, Lawrence Livermore Lab Report, No. UCRL-51288, Livermore, CA, 1972.
22. Weathersby, P.K., Survanshi, S., and Meyer, P., "Spatial sensitivity of a planar positron camera." Nuclear Instruments and Methods, Vol. 220, pp. 571-574, 1984.
23. Nicholson, P.R., Thomas, J.M., and Watson, C.R., "Characterization of PDP-11 pseudo-random number generators." Proceedings of the DEC Users Society, pp. 853-862, 1978.
24. Knuth, D.E. "The art of computer programming." Seminumerical Algorithms, Vol. II, Chap. 3, Addison-Wesley, Reading MA, 1969.
25. Snyder, D.L. and Miller, M.I., "The use of sieves to stabilize images produced with the EM algorithm for emission tomography." IEEE Transactions on Nuclear Science, Vol. NS-32, pp. 3864-3872, 1985.
26. Weathersby, P.K. and Homer, L.D., "Solubility of gases in biological fluids and tissues: a review." Undersea Biomedical Research, Vol. 7, pp. 277-296, 1980.

27. Floyd, C.E., Jaszczak, R.J., Coleman, R.E. "Convergence of the maximum likelihood reconstruction algorithm for emission computed tomography." Physics in Medicine and Biology, Vol. 32, pp. 463-476, 1987.
28. Llacer, J., Veklerov, E., and Hoffman, E.J. On the convergence of the maximum likelihood method of tomographic image reconstruction. Lawrence Berkeley Laboratory Report, No. LBL-21800, 1987 (also Proceedings of Conference on Medical Imaging, SPIE v. 767, 1987).
29. Llacer, J. and Veklerov, E., The high sensitivity of the maximum likelihood estimator method of tomographic image reconstruction. Lawrence Berkeley Laboratory Report, No. LBL-21874, 1987 (also Proceedings of Symposium on Computer Assisted Radiology (CAR'87) 1987).